Semiparametric logistic regression with unknown sizes, and its application to bioassays

Wei Zhang

Department of Statistics, University of California, Riverside, CA, 92521
wxz118@yahoo.com

Abstract

Logistic regression with unknown sizes has many important applications in biological and medical sciences. All models about this problem in the literature are parametric ones. A semiparametric regression model is proposed. This model incorporates overdispersion due to the variation of sizes, and allows general dose-response relations. An Expectation Conditional Maximization algorithm is provided to maximize the log likelihood. The bootstrap method can be used to construct confidence intervals for regression coefficients. Simulation is performed to study the behavior of the proposed model. Two real examples are investigated by the proposed model.

Keywords: Colony formation assay; Dose response; Mixture model; Quantal response

1 Introduction

Consider that there are r observations $(y_i, \boldsymbol{x}_i)$, $i = 1, 2, \ldots, r$, where y_i is a binomial random variable with size n_i and probability p_i and $\boldsymbol{x}_i$ is a vector of covariates of length ϱ. The issue of interest is to investigate how the covariates $\boldsymbol{x}_i$ affect the probabilities p_i. A logistic regression problem arises when the sizes n_i are known (e.g., McCullagh and Nelder 1999). It can happen that the sizes n_i are unknown.

The author was motivated to study the logistic regression problem with unknown sizes by colony formation assays. These assays are used to assess the cytotoxic effects of chemical or physical agents on proliferating cells. In these experiments, cells are exposed to the agent of interest, and then placed onto culture plates for colony formation. After some time, visible colonies on each plate are counted to decide how many cells survive. The initial number of cells put onto each plate is usually unknown. Table 1 presents an example in which the survival of *M. bovis* cells was studied (Trajstman 1989). Note that y_i is the number of colonies, and n_i is the unknown total number of cells on a culture plate.

There are many other applications. For example, Margolin et al. (1981) studied the effects of quinoline on the number of revertant colonies of Salmonella strain TA98. Bailer and Piegorsch (2000) reviewed the statistical methods on aquatic toxicology studies and took the effect of nitrofen on the offspring of *C. dubia* as an example. Morton (1981) presented an example of wheat disinfestation by hot air. Elder (1996) investigated the survival of V79-473 cells and their exposure times to high temperature. The radiation damage on jejunal crypts has been studied extensively (e.g., Khan et al. 1997, Kinashi et al. 1997, Mason et al. 1999, Salin et al. 2001, and Goel et al. 2003).

In the literature, the response y_i is usually assumed to be a Poisson random variable, such as Wadley (1949) and Margolin et al. (1981). Such an approximation is inappropriate when n_i and p_i are moderate in size (e.g., Elder et al. 1999). Anscombe (1949) considered overdispersion relative to the Poisson distribution and developed a model based on the negative-binomial distribution. Baker et al. (1980) treated y_i as a Poisson random variable. The y_i in the control group have a common mean m and those in the treatment group mp_i, where a probit dose-response relation is assumed. Trajstman (1989) modified the method of Baker et al. (1980) to allow a logistic dose-response relation and incorporated overdispersion by assuming a scaled Poisson variance-mean relationship. Morgan and Smith (1992) also based their work on Baker et al. (1980), and used a negative-binomial variance/mean relationship with a heterogeneity factor to handle extra Poisson variation. Kim and Taylor (1994) and Elder et al. (1999) developed

Table 1: The *M. bovis* cell survival data.

%weight/volume	No. of *M. bovis* colonies at stationarity										sample mean
	control experiment (no decontaminant)										
	52	80	55	50	58	50	43	50	53	54	51.8
	44	51	34	37	46	56	64	51	67	40	
[HPC]	decontaminant: HPC										
0.75	2	4	8	9	10	1	0	5	14	7	6.0
0.375	11	12	13	12	11	13	17	16	21	2	12.8
0.1875	16	6	20	23	23	39	18	23	33	21	22.2
0.09375	33	46	42	18	35	20	19	29	41	36	31.9
0.075	30	30	27	53	51	39	31	36	38	22	35.7
0.0075	53	62	38	54	54	38	46	58	54	57	51.4
0.00075	3	42	45	49	32	39	40	34	45	51	38.0
[Oxalic acid]	decontaminant: Oxalic acid										
5	14	15	6	13	4	1	9	6	12	13	9.3
0.5	27	33	31	30	26	41	33	40	31	20	31.2
0.05	33	26	32	24	30	52	28	28	26	22	30.1
0.005	36	54	31	37	50	73	44	50	37		45.8

a quasi-likelihood approach by regarding $y_i|n_i$ as a binomial random variable. Kim and Taylor (1994) assumed that $E(n_i) = \lambda_i$ and var $(n_i) = \lambda_i \nu$ with λ_i known and $\nu \geqslant 1$ unknown. Elder et al. (1999) estimated $\lambda = E(n_i)$ with var $(n_i) = \lambda(1 + \nu\lambda)$ and $\nu \geqslant 0$. All previous methods used parametric models.

We propose a semiparametric regression model, in which each n_i is assumed to be a Poisson random variable with mean λ_i, and the λ_i are assumed to arise as a random sample from an unspecified mixing distribution. By doing this, a rich pool of distributions can be used for λ_i.

In Section 2, a semiparametric model is formulated, and an Expectation Conditional Maximization (ECM) algorithm that maximizes the log likelihood is described. The issues of selecting the number of support points and using the bootstrap method are also discussed. Simulation results are shown in Section 3. Section 4 applies the proposed model to two real examples. One is from an *M. bovis* cell survival assay, and the other from a jejunal crypt stem cell survival assay.

2 Methods

2.1 A semiparametric model

The probability p_i can be written as $p_i = h(\boldsymbol{x}_i; \boldsymbol{\beta})$, where h is the inverse of a link function, e.g., $h^{-1} =$ logit or probit. Note that h is a general function of $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$. The unknown size n_i is assumed to be a Poisson random variable with mean λ_i. It is clear that y_i given λ_i is a Poisson random variable with mean $\lambda_i h(\boldsymbol{x}_i; \boldsymbol{\beta})$. The nuisance parameters λ_i are further assumed to follow a mixing distribution G. Because the parameter of interest $\boldsymbol{\beta}$ is in the ϱ-dimensional Euclidean space, a semiparametric regression model arises when G is treated nonparametrically. The density of a single generic observation $(y, \boldsymbol{x})$ is

$$f(y; \boldsymbol{x}, \boldsymbol{\beta}, G) = \int f(y; \boldsymbol{x}, \boldsymbol{\beta}, \lambda) dG(\lambda),$$

where $f(y;\boldsymbol{x},\boldsymbol{\beta},\lambda)$ is a Poisson density with mean $\lambda h(\boldsymbol{x};\boldsymbol{\beta})$, i.e.,

$$f(y;\boldsymbol{x},\boldsymbol{\beta},\lambda)=\exp\{-\lambda h(\boldsymbol{x};\boldsymbol{\beta})\}\{\lambda h(\boldsymbol{x};\boldsymbol{\beta})\}^y/y!,\quad y=0,1,\ldots.$$

The log likelihood can be written as

$$\ell(\boldsymbol{\beta},G)=\sum_{i=1}^{r}\log f(y_i;\boldsymbol{x}_i,\boldsymbol{\beta},G). \tag{1}$$

2.2 An ECM algorithm

In order to maximize $\ell(\boldsymbol{\beta},G)$ in (1), first we will consider the case that G is a discrete distribution with a fixed number of support points. Let $G=\sum_{j=1}^{K}\alpha_j\delta(\lambda_j)$, where $\sum_{j=1}^{K}\alpha_j=1$, $\alpha_j\geq 0$, δ is the indicator function, and $\lambda_j\in(0,\infty)$. Let $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\ldots,\alpha_K)'$, $\boldsymbol{\lambda}=(\lambda_1,\lambda_2,\ldots,\lambda_K)'$ and $\boldsymbol{\theta}$ =($\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, $\boldsymbol{\lambda}$). The log likelihood $\ell(\boldsymbol{\beta},G)$ in (1) can be written as

$$\ell(\boldsymbol{\theta})=\sum_{i=1}^{r}\log\left\{\sum_{j=1}^{K}\alpha_j f_j(y_i;\boldsymbol{x}_i,\boldsymbol{\beta},\lambda_j)\right\}. \tag{2}$$

One may consider using an EM algorithm to maximize $\ell(\boldsymbol{\theta})$ in (2). However, the M-step in the EM algorithm may be computationally unreliable.

We will consider an ECM algorithm (Meng and Rubin 1993; McLachlan and Peel 2000, p148). The ECM algorithm simplifies the M-step by replacing the complicated M-step with three computationally simpler and stabler conditional maximization (CM) steps. It also drives up the log likelihood at each iteration (Meng and Rubin 1993).

Suppose the missing datum is $\boldsymbol{z}=(z_1,z_2,...,z_K)'$, the indicator vector for the pair $(\boldsymbol{x},y)$, where $z_j=1$ for some j and $z_k=0$ for all $k\neq j$, i.e., $\lambda=\lambda_j$, $j=1,2,\ldots,K$. Note that $\boldsymbol{z}$ is multinomial distributed with size one and probability $\boldsymbol{\alpha}$. The complete density for a single datum $(\boldsymbol{x},\boldsymbol{z},y)$ is $\Pi_{j=1}^{K}\left[\alpha_j f_j(y;\boldsymbol{x},\boldsymbol{\beta},\lambda_j)\right]^{z_j}$. The joint complete log likelihood is

$$\ell_c(\boldsymbol{\theta})=\sum_{i=1}^{r}\sum_{j=1}^{K}z_{ij}\big\{\log\alpha_j+\log[f_j(y_i;\boldsymbol{x}_i,\boldsymbol{\beta},\lambda_j)]\big\}.$$

The expected conditional log likelihood to be maximized is

$$W(\boldsymbol{\theta};\boldsymbol{\theta}^{(0)})=E_{\boldsymbol{\theta}^{(0)}}\left\{\ell_c(\boldsymbol{\theta})|y_1,y_2,...,y_r\right\}.$$

The E-step involves getting the conditional expectation of z_{ij}, i.e.,

$$\pi_{ij}^{(0)}=E_{\boldsymbol{\theta}^{(0)}}\big(z_{ij}|y_1,y_2,...,y_r\big)=\frac{\alpha_j^{(0)}f_j(y_i;\boldsymbol{x}_i,\boldsymbol{\beta}^{(0)},\lambda_j^{(0)})}{\sum_{h=1}^{K}\alpha_h^{(0)}f_h(y_i;\boldsymbol{x}_i,\boldsymbol{\beta}^{(0)},\lambda_h^{(0)})}$$

for $i=1,2,\ldots,r$ and $j=1,2,\ldots,K$.

In the CM-step, we need to maximize the expected conditional complete log likelihood

$$\begin{aligned} W(\boldsymbol{\theta};\boldsymbol{\theta}^{(0)}) &= \sum_{i=1}^{r}\sum_{j=1}^{K}\pi_{ij}^{(0)}\log\alpha_j + \sum_{i=1}^{r}\sum_{j=1}^{K}\pi_{ij}^{(0)}\log f_j(y_i;\boldsymbol{x}_i,\boldsymbol{\beta},\lambda_j) \\ &= \text{constant} + \underbrace{\sum_{i=1}^{r}\sum_{j=1}^{K}\pi_{ij}^{(0)}\log\alpha_j}_{T_1(\boldsymbol{\alpha})} \\ &\quad + \underbrace{\sum_{i=1}^{r}\sum_{j=1}^{K}\pi_{ij}^{(0)}\left\{y_i\log\lambda_j + y_i\log h(\boldsymbol{x}_i;\boldsymbol{\beta}) - \lambda_j h(\boldsymbol{x}_i;\boldsymbol{\beta})\right\}}_{T_2(\boldsymbol{\beta},\boldsymbol{\lambda})} \end{aligned}$$

over $\boldsymbol{\alpha},\boldsymbol{\lambda},\boldsymbol{\beta}$ sequentially. The maximum likelihood estimator (MLE) for $\boldsymbol{\alpha}$ is

$$\alpha_j^{(1)} = r^{-1}\sum_{i=1}^{r}\pi_{ij}^{(0)},\ j = 1,2,...,K. \tag{3}$$

The conditional MLE for $\boldsymbol{\lambda}$ given $\boldsymbol{\beta} = \boldsymbol{\beta}^{(0)}$ is

$$\lambda_j^{(1)} = \frac{\sum_{i=1}^{r}\pi_{ij}^{(0)}y_i}{\sum_{i=1}^{r}\pi_{ij}^{(0)}h(\boldsymbol{x}_i;\boldsymbol{\beta}^{(0)})},\ j = 1,2,...,K. \tag{4}$$

The conditional MLE for $\boldsymbol{\beta}$ given $\boldsymbol{\lambda} = \boldsymbol{\lambda}^{(1)}$ is

$$\boldsymbol{\beta}^{(1)} = \underset{\boldsymbol{\beta}\in\mathcal{R}^{\varrho}}{\operatorname{argmax}}\, T_2(\boldsymbol{\beta},\boldsymbol{\lambda}^{(1)}). \tag{5}$$

Since there is no analytic solution for $\boldsymbol{\beta}^{(1)}$ in the optimization problem defined in (5), a Newton Raphson algorithm is applied. The first order derivative of $T_2(\boldsymbol{\beta},\boldsymbol{\lambda}^{(1)})$ is

$$\frac{\partial T_2}{\partial\boldsymbol{\beta}} = \sum_{i=1}^{r}\sum_{j=1}^{K}\pi_{ij}^{(0)}\left[\frac{y_i}{h(\boldsymbol{x}_i;\boldsymbol{\beta})} - \lambda_j^{(1)}\right]\nabla_{\boldsymbol{\beta}}h(\boldsymbol{x}_i;\boldsymbol{\beta}),$$

and the second order derivative is

$$\frac{\partial^2 T_2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'} = \sum_{i=1}^{r}\sum_{j=1}^{K}\pi_{ij}^{(0)}\left\{\left[\frac{y_i}{h(\boldsymbol{x}_i;\boldsymbol{\beta})} - \lambda_j^{(1)}\right]\frac{\partial^2 h(\boldsymbol{x}_i;\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'} - \frac{y_i\nabla_{\boldsymbol{\beta}}h(\boldsymbol{x}_i;\boldsymbol{\beta})\nabla'_{\boldsymbol{\beta}}h(\boldsymbol{x}_i;\boldsymbol{\beta})}{h(\boldsymbol{x}_i;\boldsymbol{\beta})^2}\right\}.$$

The Newton Raphson algorithm is defined by, with $\boldsymbol{\beta}_{(0)}$ $=\boldsymbol{\beta}^{(0)}$,

$$\boldsymbol{\beta}_{(t+1)} = \boldsymbol{\beta}_{(t)} - \left[\frac{\partial^2 T_2}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_{(t)}}\right]^{-1}\left[\frac{\partial T_2}{\partial\boldsymbol{\beta}}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}_{(t)}}\right]. \tag{6}$$

2.3 Selecting the number of support points

By increasing the number of support points of G, the maximized log likelihood $\ell(\hat{\boldsymbol{\theta}})$ can be increased. One may consider using the global maximizer by trying different values of K. In order to obtain a reasonable and parsimonious

Table 2: Simulation results: sd stands for standard deviation, qi for 95% quantile interval, and mse for mean square error.

setting	β_1	(α_1, α_2)	(λ_1, λ_2)	bias	sd	qi	mse
1	1	(0.5,0.5)	(100,300)	0.001	0.030	(0.942, 1.064)	0.001
2	1	(0.25,0.75)	(100,300)	0.003	0.025	(0.954, 1.049)	0.001
3	1	(0.5,0.5)	(450,650)	0.001	0.019	(0.966, 1.040)	0.000
4	1	(0.25,0.75)	(450,650)	-0.000	0.017	(0.968, 1.032)	0.000
5	2	(0.5,0.5)	(100,300)	0.007	0.072	(1.871, 2.156)	0.005
6	2	(0.25,0.75)	(100,300)	0.007	0.063	(1.901, 2.136)	0.004
7	2	(0.5,0.5)	(450,650)	-0.000	0.045	(1.919, 2.093)	0.002
8	2	(0.25,0.75)	(450,650)	0.002	0.038	(1.928, 2.076)	0.001

fit to the data, we propose to choose the number of support points by minimizing the BIC (e.g., Wang et al. 1996), i.e.,

$$\widehat{K} = \underset{K \in \{1,2,\dots\}}{\operatorname{argmin}} \{-2\ell(\hat{\boldsymbol{\theta}}) + \log(r)(2K - 1 + \varrho)\}.$$

2.4 The bootstrap method

The bootstrap method can be applied to obtain confidence intervals for the regression coefficients $\boldsymbol{\beta}$. For a random design, the nonparametric bootstrap method can be applied, in which one can sample the pairs $(y_i, \boldsymbol{x}_i)$. For a fixed design, we propose to use a parametric bootstrap method. A resample of size r is generated as follows,

$$y_i^* \sim f(y; \boldsymbol{x}_i, \widehat{\boldsymbol{\beta}}, \lambda_i), i = 1, 2, \dots, r,$$

where λ_i is a random variable drawn from the estimated mixing distribution $\widehat{G}$,

$$\widehat{G} = \sum_{j=1}^{\widehat{K}} \hat{\alpha}_j \delta(\hat{\lambda}_j).$$

3 Simulation

We report a simulation study in which there is a single covariate x. There are 10 replications for each integer x in $[-5, 5]$, so that $r = 110$. A logistic dose-response relation is assumed, i.e.,

$$\log\left\{\frac{p_i}{1 - p_i}\right\} = \beta_0 + \beta_1 x_i.$$

The intercept β_0 is fixed to be one. A 2^3 design is considered, i.e.,

$$\underbrace{\{1, 2\}}_{\beta_1} \times \underbrace{\{(0.5, 0.5), (0.25, 0.75)\}}_{(\alpha_1, \alpha_2)} \times \underbrace{\{(100, 300), (450, 650)\}}_{(\lambda_1, \lambda_2)}.$$

For each setting, 800 samples are generated. The results are shown in Table 2. One can observe that the bias, standard deviation and mean square error of the slope β_1 are quite small. The β_1 falls into the 95% quantile interval, with ends being 2.5% and 97.5% quantiles.

Table 3: The estimates of the mixing distribution and the BIC for the *M. bovis* data.

component number (j)	mixing probabilities (α_j)	support point (λ_j)	BIC
		one-component mixture	
1	1	71.98	1061.1
		two-component mixture	
1	0.048	9.601	998.0
2	0.952	73.59	
		three-component mixture	
1	0.046	9.391	$\boxed{977.0}$
2	0.840	69.52	
3	0.115	107.1	
		four-component mixture	
1	0.045	9.376	984.2
2	0.180	57.53	
3	0.697	73.84	
4	0.079	110.9	

4 Example

4.1 An *M. bovis* cell survival assay

The data in Table 1 are part of Table 1 in Trajstman (1989) and also studied by Morgan and Smith (1992). *M. bovis* cells were treated with one of the decontaminants, HPC or oxalic acid with one concentration, then placed on the culture plates for colony formation. After 12 weeks (at stationarity), the *M. bovis* colonies were counted. Trajstman (1989) and Morgan and Smith (1992) treated the count of three colonies for HPC dose at 0.00075 as an extreme observation and omitted it from all analysis. However, such a small count can be automatically taken care of in the proposed semiparametric model.

An ANOVA model is fitted with a separate factor for each level of the decontaminants. Let x_j denote a factor for the concentration level j of the decontaminants. It is assumed that the p_i satisfy that

$$\log\Big\{\frac{p_i}{1-p_i}\Big\} = \beta_0 + \sum_{j=1}^{11} \beta_j x_{ij}, \quad i = 1, 2, \ldots, 129, \tag{7}$$

where β_0 is the control effect and β_j is the effect difference between dose j and the control dose, $j = 1, 2, \ldots, 11$.

The results of estimated mixing distributions are in Table 3. The smallest BIC corresponds to $K = 3$. When $K = 3$, the estimate $\widehat{G}$ is written as

$$\widehat{G} = 0.046\,\delta(9.391) + 0.840\,\delta(69.52) + 0.115\,\delta(107.1).$$

Table 4 presents the results for the regression coefficients. In the bootstrap, 200 resamples are drawn. The bootstrap standard errors of the regression coefficients are small. Since all 95% confidence intervals except those of β_0 and β_6 do not include 0, all treatment doses except HPC 0.0075 have more negative effects on survival of *M. Bovis* cells than the control. The MLEs $\hat{\beta}_6$ and $\hat{\beta}_9$ violate the dose-response monotonicity relationship, i.e., increased negative effects on the response associated with increasing dosage of the decontaminants. This is consistent with the monotonicity violation in their sample means in Table 1. More investigation is needed for the data. The estimates $\hat{\beta}_j$ are not comparable with those in Trajstman (1989) and Morgan and Smith (1992), which used a simple linear model in (7). Figure 1 presents the responses y and their fitted values, which shows that the model fits very well.

Table 4: The estimated regression coefficients, bootstrapped standard error, and 95% confidence interval for the *M. Bovis* data.

dose	β	MLE	se	95% ci
control				
	β_0	0.882	0.117	(0.670, 1.125)
HPC				
0.75	β_1	-3.131	0.209	$(-3.615, -2.758)$
0.375	β_2	-2.317	0.188	$(-2.691, -1.946)$
0.1875	β_3	-1.639	0.180	$(-1.983, -1.293)$
0.09375	β_4	-1.294	0.176	$(-1.643, -0.960)$
0.075	β_5	-1.034	0.193	$(-1.443, -0.646)$
0.0075	β_6	0.145	0.248	$(-0.304, 0.644)$
0.00075	β_7	-0.506	0.196	$(-0.857, -0.096)$
Oxalic acid				
5	β_8	-2.715	0.184	$(-3.057, -2.363)$
0.5	β_9	-1.155	0.191	$(-1.533, -0.789)$
0.05	β_{10}	-1.251	0.182	$(-1.607, -0.874)$
0.005	β_{11}	-0.419	0.212	$(-0.807, -0.002)$

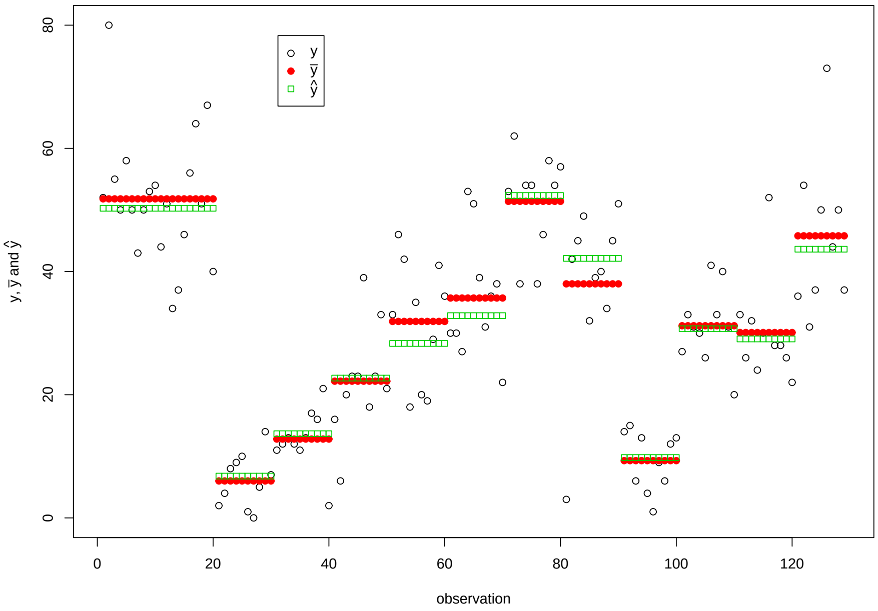

Figure 1: The response y, sample mean $\bar{y}$ and fitted value $\hat{y}$.

Table 5: Jejunal crypt data results from the proposed and previous approaches (logistic regression and Kim's method fix n_i and $E(n_i)$ at 160, respectively; Kim's and Elder's quasi-likelihood method of moments estimates come from Elder et al. (1999)).

	estimate (standard error)			
	logistic	Kim's	Elder's	proposed
β_0	7.432 (0.175)	7.410 (0.191)	6.727 (0.725)	6.705 (0.746)
β_1	−1.185 (0.024)	−1.183 (0.026)	−1.126 (0.061)	− 1.124 (0.059)
λ	—	—	194.7 (43.4)	196.1

4.2 A jejunal crypt stem cell survival assay

Table 1 in Elder et al. (1999) presents a surviving jejunal crypt data set from an experiment done on 126 mice. Note that the colony count of 12 for dose 9.25 is redundant and should be removed. Kim and Taylor (1994) also investigated the data set. A jejunal crypt is a compartment containing stem cells in a certain region of the intestine. These cells are responsible for maintaining the function of the intestine. In such an experiment, mice are treated by a certain dose of gamma rays, and then killed to count the number of surviving crypts. Because the experiment needs live mice, the total number of crypts in each mouse is unknown. It is assumed that the surviving probabilities p_i satisfy that

$$\log\left\{\frac{p_i}{1-p_i}\right\} = \beta_0 + \beta_1 x_i, \quad i = 1, 2, \ldots, 126,$$

where x_i is the gamma dose.

The BIC are 724.3 for $K = 1$ and 734.0 for $K = 2$. With $\hat{K} = 1$, the estimated $\widehat{G}$ is degenerated at $\hat{\lambda} = 196.1$. We draw 200 bootstrap resamples. Table 5 compares the estimates of the proposed method with the previous methods. All the estimates of previous methods fall into our 95% confidence intervals: $(5.089,\ 8.023)$ for β_0 and $(-1.241,\ -1.009)$ for β_1. The standard errors of the regression coefficients are quite small. Because no confidence intervals include 0, the regression coefficients are significant at the significance level of 0.05.

5 Discussion

We propose a flexible semiparametric model for the logistic regression problem with unknown sizes, in which the regression coefficients can be estimated together with the nuisance parameter, the mixing distribution.

The parameter estimates in the proposed model can be obtained effectively by an ECM algorithm. When one runs the ECM algorithm, good initial values will help find the MLEs quickly. One may run a Poisson regression analysis to find the initial values of $\boldsymbol{\beta}$.

References

[Anscombe, 1949] Anscombe, F. J. (1949). Note on a problem in probit analysis. *Annals of Applied Biology*, 36:203–205.

[Bailer and Piegorsch, 2000] Bailer, A. J. and Piegorsch, W. W. (2000). From quantal counts to mechanisms and systems: the past, present, and future of biometrics in environmental toxicology. *Biometrics*, 56:327–336.

[Baker et al., 1980] Baker, R. J., Pierce, C. B., and Pierce, J. M. (1980). Wadley's problem with controls. *GLIM Newsletter*, 3:32–35.

[Elder, 1996] Elder, J. A. (1996). *Development of quasi-likelihood techniques for the analysis of pseudo-proportional data*. Unpublished doctoral dissertation, Virginia Commonwealth University, Medical College of Virginia, Department of Biostatistics.

[Elder et al., 1999] Elder, J. A., Carter, W. H., Gennings, C., and Elswick, R. K. (1999). A quasi-likelihood approach for overdispersed binomial data when N is unobserved. *Journal of Agricultural, Biological, and Environmental Statistics*, 4:102–115.

[Goel et al., 2003] Goel, H. C., Salin, C. A., and Prakash, H. (2003). Protection of jejunal crypts by rh-3 (a preparation of hippophae rhamnoides) against lethal whole body gamma irradiation. *Phytotherapy Research*, 17:222–226.

[Khan et al., 1997] Khan, W. B., Shui, C. X., Ning, S. C., and Knox, S. J. (1997). Enhancement of murine intestinal stem cell survival after irradiation by keratinocyte growth factor. *Radiation Research*, 148(3):248–253.

[Kim and Taylor, 1994] Kim, D. K. and Taylor, J. M. G. (1994). Transform-both-sides approach for overdispersed binomial data when N is unobserved. *Journal of the American Statistical Association*, 89(427):833–845.

[Kinashi et al., 1997] Kinashi, Y., Ono, K., and Abe, M. (1997). The micronucleus assay of lymphocytes is a useful predictive assay of the radiosensitivity of normal tissue: a study of three inbred strains of mice. *Radiation Research*, 148(4):341–347.

[Margolin et al., 1981] Margolin, B. H., Kaplan, N., and Zeiger, E. (1981). Statistical analysis of the Ames salmonella/microsome test. *Proceedings of the National Academy of Sciences*, 78:3779–3783.

[Mason et al., 1999] Mason, K. A., Kishi, K., Hunter, N., Buchmiller, L., Akimoto, T., Komaki, R., and Milas, L. (1999). Effect of docetaxel on the therapeutic ratio of fractionated radiotherapy in vivo. *Clinical Cancer Research*, 5:4191–4198.

[McCullagh and Nelder, 1989] McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. Chapman and Hall, London, 2 edition.

[McLachlan and Peel, 2000] McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. Wiley.

[Meng and Rubin, 1993] Meng, X. L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika*, 80:267–278.

[Morgan and Smith, 1992] Morgan, B. J. T. and Smith, D. M. (1992). A note on Wadley's problem with overdispersion. *Applied Statistics*, 41:349–354.

[Morton, 1981] Morton, R. (1981). Generalized spearman estimators of relative dose. *Biometrics*, 37:223–233.

[Salin et al., 2001] Salin, C. A., Samanta, N., and Goel, H. C. (2001). Protection of mouse jejunum against lethal irradiation by podophylium hexandrum. *Phytomedicine*, 8(6):413–422.

[Trajstman, 1989] Trajstman, A. C. (1989). Indices for comparing decontaminants when data come from dose-response survival and contamination experiments. *Applied Statistics*, 38:481–494.

[Wadley, 1949] Wadley, F. M. (1949). Dosage-mortality correlation with number treated estimated from a parallel sample. *Annals of Applied Biology*, 36:196–202.

[Wang et al., 1996] Wang, P., Puterman, M. L., Cockburn, I., and Le, N. D. (1996). Mixed poisson regression models with covariate dependent rates. *Biometrics*, 52:381–400.